



TUGAS AKHIR - SS 141501

**PERBANDINGAN KETEPATAN KLASIFIKASI METODE REGRESI
LOGISTIK BINER DAN *MULTIVARIATE ADAPTIVE REGRESSION
SPLINES* PADA STATUS STROKE PASIEN
(Studi Kasus: RSUD Dr. H. Slamet Martodirdjo Pamekasan tahun 2015)**

**DHINTA WULANSARI TRI SURYA
NRP 1314 105 026**

Dosen Pembimbing
Ir. Sri Pingit Wulandari, M.Si
Dr. Bambang Widjanarko Otok, S.Si, M.Si

**PROGRAM STUDI S1
JURUSAN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2016**



FINAL PROJECT - SS 141501

**COMPARISON OF PRECISION CLASSIFICATION BINARY LOGISTIC
REGRESSION AND MULTIVARIATE ADAPTIVE REGRESSION SPLINES
ON TYPE OF STROKE PATIENTS
(Case Study : At Dr. H. Slamet Martodirdjo Pamekasan Hospital on 2015)**

**DHINTA WULANSARI TRI SURYA
NRP 1314 105 026**

**Supervisor
Ir. Sri Pingit Wulandari, M.Si
Dr. Bambang Widjanarko Otok, S.Si, M.Si**

**UNDERGRADUATE PROGRAMME
DEPARTMENT OF STATISTICS
FACULTY OF MATHEMATICS AND NATURAL SCIENCE
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2016**

LEMBAR PENGESAHAN

PERBANDINGAN KETEPATAN KLASIFIKASI METODE REGRESI LOGISTIK BINER DAN *MULTIVARIATE* *ADAPTIVE REGRESSION SPLINES* PADA STATUS STROKE PASIEN

(Studi Kasus : RSUD. Dr. H. Slamet Martodirdjo Pamekasan tahun 2015)

TUGAS AKHIR

Diajukan Untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana

Pada

Program Studi S1 Jurusan Statistika
Fakultas Matematika dan Ilmu Pengetahuan Alam
Institut Teknologi Sepuluh Nopember

Oleh

DHINTA WULANSARI TRI SURYA

NRP. 1314 105 026

Disetujui oleh Dosen Pembimbing Tugas Akhir

Ir. Sri Pingit Wulandari, M.Si

NIP. 19620603 198701 2 001

Disetujui oleh Dosen Co. Pembimbing Tugas Akhir :

Dr. Bambang Widjanarko Otok, S.Si, M.Si

NIP. 19681124 199412 1 001

Mengetahui
Ketua Jurusan Statistika FMIPA-ITS

Dr. Suhartono

NIP. 19710929 199512 1 001

**PERBANDINGAN KETEPATAN KLASIFIKASI METODE
REGRESI LOGISTIK BINER DAN *MULTIVARIATE ADAPTIVE
REGRESSION SPLINES*
PADA STATUS STROKE PASIEN
(Studi Kasus : RSUD. Dr. H. Slamet Martodirdjo Pamekasan tahun
2015)**

Nama Mahasiswa : Dhinta Wulansari Tri Surya
NRP : 1314 105 026
Program Studi : S1
Jurusan : Statistika FMIPA-ITS
Dosen Pembimbing : Ir. Sri Pingit Wulandari, M.Si
Dosen Co.Pembimbing : Dr. Bambang Widjanarko Otok, S.Si,
M.Si

Abstrak

Stroke merupakan penyakit yang menduduki peringkat pertama yang paling mematikan di Indonesia. Maka peneliti tertarik untuk membandingkan metode regresi logistik biner dan Multivariate Adaptive Regression Splines (MARS) pada status stroke pasien di RSUD Dr. H. Slamet Martodirdjo Pamekasan. Stroke dibedakan menjadi dua macam yaitu stroke iskemik (non hemoragik) dan hemoragik. Berdasarkan hasil ketepatan klasifikasi antara model regresi logistik biner dan MARS pada status stroke pasien di RSUD Dr. H. Slamet Martodirdjo Pamekasan pada tahun 2015 diperoleh kesimpulan bahwa model MARS memiliki nilai Total Accuracy paling tinggi dibandingkan dengan model regresi logistik biner. Sehingga pemodelan faktor yang berpengaruh terhadap status stroke pasien di RSUD Dr. H. Slamet Martodirdjo Pamekasan pada tahun 2015 lebih tepat menggunakan metode MARS. Berdasarkan hasil analisis menggunakan metode regresi logistik biner diperoleh model metode MARS diperoleh model yaitu $Y = 0,745 - 0,004 BF_1 - 0,005 BF_2 + 0,005 BF_3 + 0,003 BF_{12} - 0,010 BF_{20} + 0,008 BF_{22}$ dan model yang diperoleh dari regresi logistik biner adalah $\hat{\pi}(x) = -0,227 - 0,814 \text{ Hipertensi}(2) - 1,070 \text{ Hipertensi}(3) + 1,077 \text{ Riwayat Keluarga}(1)$.

Kata Kunci : Status Stroke, Regresi Logistik Biner, Multivariate Adaptive Regression Splines, Ketepatan Klasifikasi.

**COMPARISON OF PRECISION CLASSIFICATION BINARY
LOGISTIC REGRESSION AND MULTIVARIATE ADAPTIVE
REGRESSION SPLINES
ON TYPE OF STROKE PATIENTS
(Case Study : Dr. H. Slamet Martodirdjo Pamekasan Hospital on 2015)**

Name of Student : Dhinta Wulansari Tri Surya
NRP : 1314 105 026
Study Program : S1
Department : Statistika FMIPA-ITS
Supervisor : Ir. Sri Pingit Wulandari, M.Si
Co. Supervisor : Dr. Bambang Widjanarko Otok, S.Si, M.Si

Abstract

Stroke is one of the most deadly-caused disease in Indonesia. The researcher is interested in comparing binary logistic regression and multivariate adaptive regression splines on type of stroke patients at Dr. H. Slamet Martodirdjo Pamekasan Hospital. Stroke can be classified into stroke non-hemorrhagic and hemorrhagic. Based on the accuracy of the classification model, MARS has a total value of accuracy highest compared with binary logistic regression model, thus the case about type of stroke patients at Dr. H. Slamet Martodirdjo Pamekasan Hospital on 2015 more precise using method by MARS. While based on using MARS method obtained model is $Y = 0,745 - 0,004 BF_1 - 0,005 BF_2 + 0,005 BF_3 + 0,003 BF_{12} - 0,010 BF_{20} + 0,008 BF_{22}$ and analysis using binary logistic regression method obtain model is-

$\hat{\pi}(x) = -0,227 - 0,814 \text{ Hipertensi}(2) - 1,070 \text{ Hipertensi}(3) + 1,077 \text{ Riwayat Keluarga}(1)$.

Keywords : *Type of Stroke, Binary Logistic Regression, Multivariate Adaptive Regression Splines and Accuracy.*

DAFTAR ISI

halaman

HALAMAN JUDUL	i
LEMBAR PENGESAHAN	iv
ABSTRAK	vi
ABSTRACT	viii
KATA PENGANTAR	x
DAFTAR ISI	xii
DAFTAR TABEL	xiv
DAFTAR LAMPIRAN	xvi
BAB I PENDAHULUAN	
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah	3
1.3 Tujuan.....	3
1.4 Batasan Masalah	3
1.5 Manfaat.....	4
BAB II TINJAUAN PUSTAKA	
2.1 <i>Multivariate Adaptive Regression Spline</i> (MARS).....	5
2.2 Regresi Logistik Biner	8
2.3 Ketepatan Klasifikasi.....	13
2.4 Pengertian Penyakit Stroke	14
2.5 Klasifikasi Stroke.....	14
2.6 Faktor Penyebab Stroke	15
BAB III METODOLOGI PENELITIAN	
3.1 Sumber Data.....	19
3.2 Variabel Penelitian.....	21
3.3 Langkah Analisis Data	21
BAB IV ANALISIS DAN PEMBAHASAN	
4.1 Statistika Deskriptif Faktor-faktor yang Berpengaruh terhadap Status Stroke Pasien RSUD Dr. H. Slamet Martodirdjo Pamekasan tahun 2015	23
4.2 <i>Multivariate Adaptive Regression Spline</i> (MARS)	24
4.3 Analisis Regresi Logistik Biner.....	28
4.4 Ketepatan Klasifikasi.....	31
BAB V KESIMPULAN DAN SARAN	
5.1 Kesimpulan	35
5.2 Saran	35
DAFTAR PUSTAKA	37

LAMPIRAN41
BIODATA PENULIS49

DAFTAR TABEL

halaman

Tabel 3.1 Variabel Faktor-faktor yang Mempengaruhi Status Stroke	19
Tabel 4.1 Hasil Analisis Statistika Deskriptif.....	23
Tabel 4.2 Kombinasi BF, MI dan MO.....	24
Tabel 4.3 Nilai <i>Odds Ratio</i> untuk Model MARS	26
Tabel 4.4 Tingkat Kepentingan Variabel dalam MARS	28
Tabel 4.5 Uji Signifikan Parameter secara Univariabel.....	29
Tabel 4.6 Uji Signifikan Parameter secara Serentak.....	29
Tabel 4.7 Uji Signifikan Parameter secara Parsial.....	30
Tabel 4.8 Hasil Klasifikasi Model Regresi Logistik Biner	32
Tabel 4.9 Hasil Klasifikasi Model MARS	32
Tabel 4.10 Hasil Perhitungan Nilai <i>Sensitivity</i> , <i>Specificity</i> dan <i>Total Accuracy</i>	32

DAFTAR LAMPIRAN

halaman

Lampiran A.	Data Hasil Rekam Medik di RSUD Dr. H. Slamet Martodirdjo Pamekasan tahun 2015.....	41
Lampiran B.	Hasil <i>Output</i> SPSS	42
Lampiran C.	Hasil <i>Output</i> MARS 2.0.....	47

BAB I

PENDAHULUAN

1.1 Latar Belakang

Stroke merupakan penyakit yang menduduki peringkat pertama yang paling mematikan di Indonesia (Widowati, 2015). Sedangkan menurut Dinas Kesehatan kabupaten Pamekasan, penyakit pada sistem otot dan jaringan termasuk stroke menjadi peringkat pertama pada tahun 2012 yaitu sebanyak 87.456 kasus (Dinkes, 2012). Hal ini tentu menjadi pusat perhatian bagi pengamat kesehatan sehingga penulis tertarik untuk mengangkat kasus ini sebagai topik penelitian. Stroke adalah serangan yang terjadi pada otak secara mendadak akibat tersumbat atau pecahnya pembuluh darah otak (Dr. Gaul, 2013). Dr. Gaul mengatakan bahwa dengan kata lain, penyakit stroke merupakan penyakit pembuluh darah otak (serebrovaskuler) yang ditandai dengan kematian jaringan otak (infark serebral) dipicu oleh penyumbatan atau penyempitan dan pecahnya pembuluh darah yang menuju ke otak, sehingga pasokan darah dan oksigen yang menuju ke otak mengalami kekurangan maka serangkaian reaksi biokimia dapat merusak dan mematikan sel-sel saraf otak.

Stroke dibedakan menjadi dua macam yaitu stroke iskemik dan hemoragik (Batticaca, 2012). Menurut Batticaca, stroke iskemik merupakan komplikasi dari penyakit vaskular, yang ditandai dengan gejala penurunan tekanan darah yang mendadak, takikardia, pucat dan pernafasan yang tidak teratur dan stroke hemoragik pada umumnya disebabkan oleh adanya pendarahan intrakranial dengan gejala peningkatan tekanan darah sistole lebih dari 200 mmHg pada hipertoni dan 180 mmHg pada normotoni, bradikardia, wajah keunguan, sianosis dan pernafasan mengorok. Faktor-faktor yang menyebabkan stroke yaitu meliputi hipertensi, hipotensi, obesitas, kolesterol darah tinggi, riwayat penyakit jantung, riwayat penyakit diabetes melitus, merokok dan stres (Batticaca, 2012).

Berbagai penelitian sebelumnya yang berkaitan dengan stroke, yaitu Oktafia (2013) melakukan penelitian mengenai faktor-faktor yang mempengaruhi laju perbaikan kondisi klinis pasien penderita stroke dengan regresi *cox weibull* dengan hasil yang diperoleh yaitu faktor-faktor yang berpengaruh terhadap laju perbaikan klinis pasien stroke di RSUD Haji Surabaya pada periode Januari hingga Agustus pada tahun 2012 adalah usia, penyakit jantung, diabetes melitus, hiperkolesterol, *TIA* dan jenis stroke. Christanto, Mahama dan Tumboimbela (2014) meneliti tentang faktor-faktor risiko pada pasien stroke yang dirawat inap di Irina F Neurologi RSUP Prof. Dr. R. D. Kandou Manado periode Januari – Oktober 2012 dengan hasil bahwa terdapat tiga faktor risiko yang berpengaruh terhadap kejadian stroke yaitu umur, jenis kelamin dan hipertensi. Surya (2014) menganalisis mengenai faktor-faktor yang mempengaruhi status penyakit stroke pasien di RSUD Dr. H. Slamet Martodirdjo Pamekasan dengan menggunakan Metode Regresi Logistik Biner diperoleh hasil bahwa variabel yang signifikan terhadap status stroke adalah pasien dengan jenis kelamin laki-laki, kolesterol tinggi dan penderita asam urat. Kabi, Tumewah, Kembuan (2015) melakukan penelitian mengenai gambaran faktor risiko pada penderita stroke iskemik yang dirawat inap neurologi RSUP Prof. Dr. R. D. Kandou Manado Periode Juli 2012 – Juni 2013 diperoleh hasil bahwa pasien yang sering terkena stroke adalah pasien yang berumur 51 – 65 tahun dan pasien yang memiliki riwayat hipertensi. Namun, penelitian-penelitian tersebut belum ada yang membandingkan hasil ketepatan analisis menggunakan metode regresi logistik biner dan MARS pada penyakit stroke. Oleh karena itu, peneliti tertarik untuk membandingkan kedua metode tersebut pada status stroke pasien di RSUD Dr. H. Slamet Martodirdjo Pamekasan.

Model regresi logistik biner merupakan salah satu model regresi logistik yang digunakan untuk menganalisa hubungan antara satu variabel respon dan beberapa variabel prediktor, dimana variabel respon terdiri dari data kualitatif yang bersifat

dikotonomi. Sedangkan metode MARS merupakan metode yang tidak bergantung pada asumsi bentuk kurva tertentu, sehingga memberikan fleksibilitas yang lebih besar dalam penggunaannya (Friedman, 1991). Penelitian mengenai model klasifikasi menggunakan metode regresi logistik dan *multivairate adaptive regression* (MARS) dilakukan oleh Sumarno (2014) dimana studi kasus dilakukan pada data survei biaya hidup (SBH) Kota Kediri tahun 2012 dengan hasil analisis yaitu metode MARS mempunyai ketepatan klasifikasi lebih baik jika dibandingkan dengan regresi logistik. Dalam penelitian ini membahas mengenai perbandingan ketepatan klasifikasi metode regresi logistik biner dan MARS pada status stroke yang diderita oleh pasien di RSUD Dr. H. Slamet Martodirdjo Pamekasan tahun 2015.

1.2 Rumusan Masalah

Berdasarkan pada latar belakang diatas, maka permasalahan penelitian ini adalah sebagai berikut.

1. Bagaimana pemodelan faktor yang mempengaruhi status stroke pasien di RSUD Dr. H. Slamet Martodirdjo Pamekasan tahun 2015 dengan metode regresi logistik biner dan MARS?
2. Bagaimana ketepatan klasifikasi mengenai status stroke pasien di RSUD Dr. H. Slamet Martodirdjo Pamekasan tahun 2015?

1.3 Tujuan

Tujuan dari penelitian ini adalah sebagai berikut.

1. Memodelkan faktor yang mempengaruhi status stroke pasien di RSUD Dr. H. Slamet Martodirdjo Pamekasan tahun 2015 dengan metode regresi logistik biner dan MARS.
2. Mendapatkan metode yang lebih tepat untuk dipergunakan pada klasifikasi status stroke pasien di RSUD Dr. H. Slamet Martodirdjo Pamekasan tahun 2015.

1.4 Batasan Masalah

Batasan masalah yang digunakan pada penelitian ini adalah pasien rawat inap yang menderita stroke di RSUD Dr. H. Slamet Martodirdjo Pamekasan pada tahun 2015. Status stroke pasien ditentukan ketika saat masuk ke rumahsakit.

1.5 Manfaat

Manfaat yang akan diperoleh pada penelitian ini sebagai berikut.

1. Memberikan wawasan keilmuan yang lebih kepada penulis khususnya dan kepada masyarakat pada umumnya tentang penggunaan metode regresi logistik biner dan MARS.
2. Mendapatkan informasi tentang faktor yang mempengaruhi status stroke pasien di RSUD Dr. H. Slamet Martodirdjo Pamekasan pada tahun 2015.

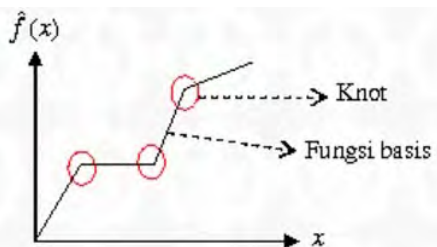
BAB II TINJAUAN PUSTAKA

2.1 *Multivariate Adaptive Regression Spline (MARS)*

Friedman (1991) mengartikan MARS yaitu pendekatan untuk regresi multivariat nonparametrik yang difokuskan untuk mengatasi permasalahan dimensi yang tinggi dan diskontinuitas pada data, sehingga menghasilkan prediksi variabel respon yang akurat. Metode MARS tidak memerlukan asumsi tentang hubungan fungsional yang mendasar antara variabel respon dan prediktor. Selain itu, MARS merupakan pengembangan dari pendekatan *Recursive Partition Regression (RPR)* yang masih memiliki kelemahan dimana model yang dihasilkan tidak kontinu pada titik knot (Friedman, 1991).

Menurut Nash dan Bradford (2001) menyatakan bahwa ada beberapa hal yang perlu diperhatikan dalam membangun model MARS, yaitu:

1. Knot, yaitu akhir sebuah garis regresi (region) dan awal sebuah garis (region) yang lain. Di setiap titik knot, diharapkan adanya kontinuitas dan fungsi basis antar satu region dengan region lainnya. Sebagai ilustrasi, suatu MARS spline dengan tiga knot digambarkan seperti Gambar 2.1.



Gambar 2.1 Ilustrasi MARS Spline dengan Tiga Knot.

2. Fungsi Basis, yaitu suatu fungsi yang digunakan untuk menjelaskan hubungan antara variabel respon dan variabel prediktor. Fungsi basis ini merupakan fungsi parametrik yang didefinisikan pada tiap regresi dan biasa disebut

sebagai *truncated power*. Fungsi basis yang dipilih pada umumnya berbentuk polinomial dengan turunan yang kontinu pada setiap knot. Menurut Cherkassky dan Mulier (1998) fungsi basis dapat diartikan sebagai sekumpulan fungsi yang digunakan untuk merepresentasikan informasi yang terdiri atas satu atau lebih variabel termasuk interaksi antar variabel. Satu fungsi basis adalah jarak antar knot yang berurutan.

Friedman (1991) memberi saran untuk jumlah maksimum basis fungsi (BF) adalah dua sampai empat kali jumlah variabel prediktor. Jumlah interaksi maksimum (MI) yaitu satu, dua, atau tiga dengan pertimbangan jika lebih dari tiga akan mendapatkan model yang sangat kompleks. Untuk observasi minimum (MO) antar *knot* sebanyak nol, satu, dua, tiga, lima atau sepuluh.

Model umum persamaan MARS dirumuskan oleh Friedman (1991) seperti disajikan persamaan 2.8 berikut,

$$\hat{f}(x) = a_0 + \sum_m^M a_m \prod_{k=1}^{K_m} [S_{km}(x_{v(k,m)} - t_{km})] \quad (2.1)$$

dengan:

a_0 = parameter fungsi basis induk

a_m = parameter dari fungsi basis ke- m

M = maksimum fungsi basis (*nonconstant* fungsi basis)

K_m = derajat interaksi

S_{km} = bernilai +1 jika data berada disebelah kanan knot dan bernilai -1 jika data berada disebelah kiri knot

$x_{v(k,m)}$ = variabel prediktor

t_{km} = nilai knot dari variabel prediktor $x_{v(k,m)}$

Berdasarkan persamaan (2.1), model MARS dapat dituliskan sebagai berikut.

$$\hat{f}(x) = a_0 + \sum_{m=1}^M a_m B_m(x) + \varepsilon \quad (2.2)$$

$$\text{dengan } B_m(x) = \prod_{k=1}^{K_m} [S_{km}(x_{v(k,m)} - t_{km})]$$

Dalam bentuk matriks dapat ditulis menjadi :

$$\hat{f}(x) = \mathbf{B}a + \varepsilon$$

dimana

$$y = (y_1, \dots, y_n)^T, \quad a = (a_0, \dots, a_m)^T, \quad \varepsilon = (\varepsilon_1, \dots, \varepsilon_n)^T$$

$$B = \begin{bmatrix} 1 & \prod_{k=1}^{K_1} [S_{1m}(x_{1(1,m)} - t_{1m})] & \cdots & \prod_{k=1}^{K_M} [S_{Mm}(x_{1(M,m)} - t_{Mm})] \\ 1 & \prod_{k=1}^{K_1} [S_{1m}(x_{2(1,m)} - t_{1m})] & \cdots & \prod_{k=1}^{K_M} [S_{Mm}(x_{2(M,m)} - t_{Mm})] \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \prod_{k=1}^{K_1} [S_{1m}(x_{n(1,m)} - t_{1m})] & \cdots & \prod_{k=1}^{K_M} [S_{Mm}(x_{n(M,m)} - t_{Mm})] \end{bmatrix}$$

Persamaan (2.1) dapat dijabarkan sebagai berikut.

$$\begin{aligned} \hat{f}(x) &= a_0 + \sum_m^M a_m \prod_{k=1}^{K_m} [S_{km}(x_{v(k,m)} - t_{km})] \\ \hat{f}(x) &= a_0 + \sum_m^M a_m \prod_{k=1}^{K_m} [S_{1m}(x_{v(1,m)} - t_{1m})] \\ &\quad + \sum_m^M a_m \prod_{k=1}^{K_m} [S_{1m}(x_{v(1,m)} - t_{1m})][S_{2m}(x_{v(2,m)} - t_{2m})] \\ &\quad + \sum_m^M a_m \prod_{k=1}^{K_m} [S_{1m}(x_{v(1,m)} - t_{1m})][S_{2m}(x_{v(2,m)} - t_{2m})] \end{aligned}$$

$$[S_{3m}(x_{v(3,m)} - t_{3m})] + \dots \quad (2.3)$$

Secara umum, persamaan (2.3) dapat dituliskan sebagai berikut.

$$\hat{f}(x) = a_0 + \sum_{Km=1} f_i(x_i) + \sum_{Km=2} f_{ij}(x_i, x_j) + \sum_{Km=3} f_{ijk}(x_i, x_j, x_k) + \dots \quad (2.4)$$

Pada persamaan (2.4) menunjukkan bahwa suku pertama meliputi semua fungsi basis untuk satu variabel, suku kedua meliputi semua fungsi basis untuk interaksi antara dua variabel, suku ketiga meliputi semua fungsi basis untuk interaksi antara tiga variabel dan seterusnya (Friedman, 1991).

2.2 Regresi Logistik Biner

Regresi logistik biner merupakan suatu metode analisis data yang digunakan untuk mencari hubungan antara variabel dependen (y) yang bersifat biner atau dikotomis dengan variabel independen (x) yang bersifat polikotomis (Hosmer dan Lemeshow, 2000). *Outcome* dari variabel respon y terdiri dari 2 kategori yaitu “sukses” dan “gagal” yang dinotasikan dengan $y = 1$ (sukses) dan $y = 0$ (gagal). Dalam keadaan demikian, variabel y mengikuti distribusi Bernoulli untuk setiap observasi tunggal. Fungsi distribusi Bernoulli untuk setiap observasi adalah diberikan sebagai berikut,

$$f(y) = \pi^y (1 - \pi)^{1-y}, \quad y = 0, 1 \quad (2.5)$$

dimana $y = 0$ maka $f(y) = 1 - \pi$ dan jika $y = 1$ maka $f(y) = \pi$. Nilai fungsi regresi logistiknya terletak antara 0 hingga 1. Berikut bentuk model regresi logistiknya (Hosmer dan Lemeshow),

$$\pi(x) = \frac{e^{\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p}}{1 + e^{\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p}} \quad (2.6)$$

dimana p = banyaknya variabel independen.

2.2.1 Pendugaan Parameter

Metode MLE digunakan untuk mengestimasi parameter-parameter dalam regresi logistik dan pada dasarnya metode

maksimum likelihood memberikan nilai estimasi β dengan memaksimumkan fungsi likelihoodnya (Hosmer dan Lemeshow, 2000). Secara matematis fungsi likelihood dapat dinyatakan :

$$l(\beta) = \prod_{i=1}^n \pi(x_i)^{y_i} (1 - \pi(x_i))^{1-y_i} \quad (2.7)$$

Dimana

$$\pi(x_i) = \frac{\exp(\sum_{j=0}^p \beta_j x_{ij})}{1 + \exp(\sum_{j=0}^p \beta_j x_{ij})}$$

$$\begin{aligned} L(\beta) &= \ln[l(\beta)] \\ &= \sum_{j=0}^p \left[\sum_{i=1}^n y_i x_{ij} \right] \beta_j - \sum_{i=1}^n \ln \left[1 + \exp \left(\sum_{j=0}^p \beta_j x_{ij} \right) \right] \end{aligned} \quad (2.8)$$

Persamaan tersebut diturunkan terhadap β , maka diperoleh persamaan sebagai berikut.

$$\frac{\partial L(\beta)}{\partial \beta_j} = \sum_{i=1}^n y_i x_{ij} - \sum_{i=1}^n x_{ij} \pi(x_i) \quad (2.9)$$

Setelah persamaan tersebut diturunkan terhadap β , persamaan di atas kemudian disamakan dengan nol, namun menurut (Agresti, 2002) cara tersebut sering diperoleh hasil yang eksplisit sehingga dilakukan metode iterasi Newton Rhapson untuk memaksimumkan fungsi likelihood. Algoritma iterasi Newton Rhapson dapat dituliskan sebagai berikut.

1. Menentukan nilai taksiran awal parameter $\hat{\beta}^{(0)}$ yang diperoleh dari metode *Ordinary Least Square* (OLS), yaitu $\hat{\beta}^{(0)} = (X'X)^{-1}X'Y$

dengan $X = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1p} \\ 1 & x_{21} & x_{22} & \dots & x_{2p} \\ & \vdots & & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{np} \end{bmatrix}$ dan $Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$

2. Membentuk vektor gradien g

$$g^{(t)}(\beta^{(t)}) = \left(\frac{\partial L(\beta)}{\partial \beta_0}, \frac{\partial L(\beta)}{\partial \beta_1}, \dots, \frac{\partial L(\beta)}{\partial \beta_p} \right)$$

dengan p adalah banyaknya variabel prediktor

3. Membentuk matriks Hessian H

$$H^{(t)}(\beta^{(t)}) = \begin{bmatrix} \frac{\partial^2 \ln L(\beta)}{\partial \beta_0^2} & \frac{\partial^2 \ln L(\beta)}{\partial \beta_0 \partial \beta_1} & \dots & \frac{\partial^2 \ln L(\beta)}{\partial \beta_0 \partial \beta_p} \\ \frac{\partial^2 \ln L(\beta)}{\partial \beta_0 \partial \beta_1} & \frac{\partial^2 \ln L(\beta)}{\partial \beta_1^2} & \dots & \frac{\partial^2 \ln L(\beta)}{\partial \beta_1 \partial \beta_p} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 \ln L(\beta)}{\partial \beta_0 \partial \beta_p} & \frac{\partial^2 \ln L(\beta)}{\partial \beta_1 \partial \beta_p} & \dots & \frac{\partial^2 \ln L(\beta)}{\partial \beta_p^2} \end{bmatrix}$$

4. Memasukkan nilai $\tilde{\beta}^{(0)}$ ke vektor g dan matriks H sehingga diperoleh vektor $g^{(t)}(\tilde{\beta}^{(0)})$ dan matriks $H^{(t)}(\tilde{\beta}^{(0)})$

5. Mulai dari $t = 0$ dilakukan iterasi pada persamaan

$$\beta^{(t+1)} = \beta^{(t)} - \left(H^{(t)}(\beta^{(t)}) \right)^{-1} g^{(t)}(\beta^{(t)}). \text{ Nilai } \beta^{(t)} \text{ adalah}$$

estimasi parameter yang konvergen pada iterasi ke- t

6. Apabila belum didapatkan penaksir parameter yang konvergen, maka kembali pada langkah (5) hingga iterasi ke $t = t + 1$. Iterasi berhenti jika $\|\beta^{(t+1)} - \beta^{(t)}\| \leq \varepsilon$. Hasil penaksiran yang diperoleh adalah $\beta^{(t+1)}$ pada iterasi terakhir.

2.2.2 Pengujian Parameter Model Regresi Logistik

Pengujian parameter dilakukan untuk menguji variabel prediktor berpengaruh atau tidak terhadap variabel respon. Adapun pengujian parameter yang dilakukan adalah uji signifikansi secara serentak dan uji signifikansi secara individu (Hosmer dan Lemeshow, 2000).

- a. Uji Signifikansi Parameter secara Serentak
Hipotesis yang digunakan adalah sebagai berikut.

$$H_0: \beta_1 = \beta_2 = \dots = \beta_p = 0$$

$$H_1: \text{Paling tidak terdapat satu } \beta_j \neq 0; j = 1, 2, \dots, p$$

$$\text{Statistik Uji : } G = -2 \ln \frac{\left(\frac{n_1}{n}\right)^{n_1} \left(\frac{n_0}{n}\right)^{n_0}}{\sum_{i=1}^n \hat{\pi}^{y_i} (1 - \hat{\pi}_1)^{(1-y_i)}} \quad (2.10)$$

$$\text{dimana } n_1 = \sum_{i=1}^n y_i, \quad n_0 = \sum_{i=1}^n (1 - y_i), \quad n = n_1 + n_0$$

Daerah kritis : Tolak H_0 jika nilai $G > \chi^2_{(\alpha, \text{df})}$ dengan df (*degree of freedom*) adalah banyaknya parameter dalam model tanpa β_0 .

- b. Uji Signifikansi Parameter secara Parsial
Hipotesis pengujian ini adalah sebagai berikut.

$$H_0: \beta_j = 0$$

$$H_1: \beta_j \neq 0; j = 1, 2, \dots, p$$

$$\text{Statistik Uji : } W = \frac{\beta_j}{\hat{SE}(\beta_j)} \quad (2.11)$$

Statistik uji W tersebut juga disebut sebagai statistik uji Wald, dimana mengikuti distribusi normal sehingga H_0 ditolak jika nilai $|W| > Z_{\alpha/2}$ dan dapat diperoleh melalui persamaan berikut,

$$W^2 = \frac{\beta_j^2}{\hat{SE}(\beta_j)^2} \quad (2.12)$$

Statistik uji tersebut mengikuti distribusi *Chi-Square* sehingga H_0 ditolak jika nilai $W^2 > \chi^2_{(\text{df}, \alpha)}$ dengan df = 1.

2.2.3 Uji Kesesuaian Model

Pengujian ini dilakukan untuk mengetahui apakah terdapat perbedaan yang nyata antara hasil observasi dengan prediksi

model. Pengujian kesesuaian model dilakukan dengan *Hosmer-Lemeshow Goodness-of-fit test* dengan hipotesis sebagai berikut.

H_0 : Ada hubungan antara hasil pengamatan dengan hasil prediksi

H_1 : Tidak ada hubungan antara hasil pengamatan dengan hasil prediksi

Statistik Uji:

$$\hat{C} = \sum_{k=1}^g \frac{(o_k - n'_k \bar{\pi}_k)^2}{n'_k \bar{\pi}_k (1 - \bar{\pi}_k)}$$

$$\text{Dengan } o_k = \sum_{i=1}^{C_k} y \quad \text{dan} \quad \bar{\pi}_k = \sum_{i=1}^{C_k} \frac{m_j \pi_j}{n'_k}$$

Keterangan :

g : jumlah grup (kombinasi kategori dalam model serentak)

n'_k : banyaknya observasi pada grup ke- k

O_k : observasi pada grup ke- k

$\bar{\pi}_k$: rata-rata taksiran pelang

m_j : banyaknya observasi yang memiliki nilai $\hat{\pi}_j$

Daerah Kritis: Tolak H_0 jika $\hat{C} > \chi^2_{(db, \alpha)}$ dengan derajat bebas = $k - 1$.

2.2.4 Intepretasi Koefisien Parameter

Salah satu ukuran yang digunakan untuk menginterpretasi koefisien variabel prediktor disebut *Odds ratio*. *Odds ratio* merupakan perbandingan peluang munculnya suatu kejadian dengan peluang tidak munculnya kejadian tersebut. Menurut (Hosmer and Lemeshow, 2000) *Odds ratio* didefinisikan sebagai berikut.

$$OR = \frac{\pi(1)/[1 - \pi(1)]}{\pi(0)/[1 - \pi(0)]} = \frac{\pi(1)[1 - \pi(0)]}{\pi(0)[1 - \pi(1)]} = \frac{e^{\beta_0 + \beta_1}}{e^{\beta_0}} = e^{\beta_1} \quad (2.13)$$

$$\ln(OR) = \ln \left[\frac{\pi(1)/[1-\pi(1)]}{\pi(0)/[1-\pi(0)]} \right] = g(1) - g(0) = \ln(e^{\beta_1}) = \beta_1 \quad (2.14)$$

2.3 Ketepatan Klasifikasi

Zambrano, Rocco, Claudio dan Museli (2005) mengemukakan bahwa ada beberapa ukuran yang bisa digunakan dalam mengklasifikasi yaitu ukuran *Sensitivity*, *Specificity*, *Accuracy*, *Noice Signal Ratio* (NSR), *Geometric Mean* (*G-Mean*). NSR digunakan oleh Zambrano, *et al.* (2005) untuk mengukur performa klasifikasi dari logit, probit, *Support Vector Machine* (SVM) dan *Humming Clustering* (HC) dalam mengestimasi angkatan kerja wanita. *Sensitivity* menunjukkan nilai persentase kebenaran dari kejadian-kejadian yang diinginkan terletak pada kelompok nol. Sedangkan *Specificity* menunjukkan nilai persentase kebenaran dari kejadian-kejadian yang tidak diinginkan, misalkan kejadian yang tidak diinginkan terletak pada kelompok 1. *Total Accuracy* merupakan nilai persentase ketepatan klasifikasi secara keseluruhan, nilai dari persentase NSR dipakai untuk mengukur rasio sinyal palsu (*1-sensitivity*) terhadap sinyal baik (*specificity*) yang dihasilkan pada tabel klasifikasi (Zambrano, *et al.*, 2005), sehingga semakin kecil nilai rasio dari persentase NSR maka klasifikasi yang dihasilkan adalah semakin baik, sedangkan nilai dari *G-Mean* dipakai untuk memaksimalkan akurasi dengan mempertimbangkan keseimbangan pada masing-masing kelas, sehingga semakin besar nilai *G-Mean* maka semakin baik akurasinya.

Berikut contoh tabel dua arah yang merupakan hasil klasifikasi.

Tabel 2.1 Tabel Klasifikasi Dua Arah

Kelompok Observasi	Kelompok Prediksi		
	0	1	Total
0	A	B	A + B
1	C	D	C + D
Total	A + C	B + D	A + B + C + D

Sumber : Bishop, Flenberg, dan Holland. 2007

Berikut rumus *Sensitivity*, *Specificity*, *Total Accuracy* (Bishop, *et al.*,2007),

$$Sensitivity = a^+ = \frac{D}{C + D} \quad (2.15)$$

$$Specificity = a^- = \frac{A}{A + B} \quad (2.16)$$

$$Total Accuracy = \frac{A + D}{A + B + C + D} \quad (2.17)$$

2.4 Pengertian Penyakit Stroke

Menurut Muttaqin (2008) stroke atau gangguan peredaran darah otak (GPDO) merupakan penyakit neurologis yang sering dijumpai secara cepat dan tepat. Stroke merupakan kelainan fungsi otak yang timbul mendadak yang disebabkan karena terjadinya gangguan peredaran darah otak dan bisa terjadi pada siapa saja dan kapan saja. Menurut WHO stroke adalah adanya tanda-tanda klinik yang berkembang cepat akibat gangguan fungsi otak fokal (global) dengan gejala-gejala yang berlangsung selama 24 jam atau lebih yang menyebabkan kematian tanpa adanya penyebab yang jelas selain vaskular. Stroke merupakan penyakit yang paling sering menyebabkan cacat berupa kelumpuhan anggota gerak, gangguan bicara, proses berpikir, daya ingat, dan bentuk-bentuk kecacatan yang lain sebagai akibat gangguan fungsi otak.

2.5

lasifikasi Stroke

Menurut Muttaqin (2008) stroke dibagi menjadi dua macam yaitu stroke hemoragik dan stroke nonhemoragik. Stroke hemoragik merupakan pendarahan serebral dan mungkin perdarahan subaraknoid. Disebabkan oleh pecahnya pembuluh darah otak pada area otak tertentu. Bisa terjadi saat beraktivitas maupun saat istirahat. Kesadaran klien umumnya menurun. Stroke Nonhemoragik (iskemik) dapat berupa iskemia atau

emboli dan trombosis serebral, biasanya terjadi saat setelah lama beristirahat, baru bangun tidur atau di pagi hari. Tidak terjadi pendarahan namun terjadi iskemia yang menimbulkan hipoksia dan selanjutnya dapat timbul edema sekunder. Kesadaran umumnya baik.

2.6 Faktor Penyebab Stroke

Menurut Muttaqin (2008), adapun faktor resiko stroke yaitu hipertensi, penyakit kardiovaskular, kolesterol tinggi, obesitas, peningkatan hematokrit, diabetes, kontrasepsi oral, merokok, penyalahgunaan obat dan konsumsi alkohol. Sedangkan menurut Muhlisin (2015), disebutkan bahwa adapun penyakit yang dapat menyebabkan stroke adalah tekanan darah tinggi, kolesterol tinggi, penyakit jantung, diabetes, obesitas, serangan stroke sebelumnya atau *transient ischemic (TIA)*, penyakit sel sabit, merokok, minum alkohol, kurang aktifitas fisik, riwayat keluarga, usia dan jenis kelamin.

Menurut Sulistyawati (2013) usia, jenis kelamin dan riwayat keluarga termasuk dalam kategori faktor resiko tak terkendali. Sedangkan hipertensi, penyakit jantung, diabetes, kolesterol dan merokok termasuk dalam faktor resiko terkendali. Berdasarkan literatur tersebut, maka penelitian ini menyimpulkan bahwa variabel yang akan menjadi faktor yang mempengaruhi status stroke pada pasien yaitu usia, jenis kelamin, riwayat keluarga, hipertensi, kolesterol, pekerjaan, diabetes, asam urat, berat badan dan riwayat keluarga stroke. Berikut definisi mengenai variabel-variabel tersebut.

a. Usia

Semakin bertambah usia seseorang, maka semakin beresiko. Setelah berusia 55 tahun, resikonya berlipat ganda setiap kurun waktu sepuluh tahun. Dua pertiga dari semua serangan stroke terjadi pada orang yang berusia di atas 65 tahun. Tetapi, itu tidak berarti bahwa stroke hanya terjadi pada orang lanjut usia karena stroke dapat menyerang semua kelompok umur (Sulistyawati, 2013).

b. Jenis Kelamin

Pria lebih beresiko terkena stroke daripada wanita, tetapi penelitian menyimpulkan bahwa justru lebih banyak wanita yang meninggal karena stroke. Resiko stroke pria 1,25 lebih tinggi daripada wanita, tetapi serangan stroke pada pria terjadi di usia lebih muda sehingga tingkat kelangsungan hidup juga lebih tinggi. Dengan perkataan lain, walau lebih jarang terkena stroke, pada umumnya wanita terserang pada usia lebih tua, sehingga kemungkinan meninggal lebih besar (Sulistiyawati, 2013).

c. Riwayat Keluarga

Kompas (2011) memaparkan bahwa menurut para ahli, jika salah satu orang tua anda memiliki tekanan darah tinggi, risiko anda mendapatkan penyakit ini sebesar 15% atau bahkan lebih tinggi. Sedangkan untuk kolesterol tinggi memiliki peluang diwariskan sebesar 50% (Kompas, 2011). Sejarah keluarga memegang peranan penting dalam kondisi kesehatan seseorang.

d. Hipertensi

Tekanan darah tinggi, juga disebut hipertensi, dapat sangat meningkatkan resiko stroke. Hal ini disebabkan karena tekanan darah yang terlalu tinggi dapat membuat pembuluh darah di otak menjadi pecah lalu terjadilah stroke hemoragik (Muhlisin, 2015). Menurut WHO, tekanan darah normal untuk orang dewasa adalah 120/80 mmHg, namun bila tekanan sistolik antara 120-139 dan diastolik antara 80-89 maka itu juga masih dikatakan tekanan darah normal. WHO mengklasifikan penderita hipertensi yaitu hipertensi tingkat 1 (sistolik : 140-159 mmHg dan diastolik : 90-99 mmHg), hipertensi tingkat 2 (sistolik : 160-179 mmHg dan diastolik : 100-109 mmHg) dan hipertensi tingkat 3 (sistolik : lebih dari 180 mmHg dan diastolik : lebih dari 110 mmHg).

e. Kolesterol Tinggi

Kolesterol tinggi dapat menjadi faktor resiko stroke karena kolesterol tinggi dalam darah dapat membangun timbunan lemak (plak) pada dinding pembuluh darah. Deposit lemak (plak) tersebut dapat memblokir aliran darah ke otak, menyebabkan

stroke iskemik (Muhlisin, 2015). Kadar kolesterol bisa dikatakan normal jika ukurannya kurang dari 200 mg/dL, dikatakan normal cenderung tinggi jika ukurannya 200-239 mg/dL, dan dikatakan tinggi jika lebih dari 240 mg/dL (Futurebeauty,2012).

f. Pekerjaan

Menurut Setiawan (2014), pekerjaan dalam arti luas adalah aktivitas utama yang dilakukan oleh manusia. Dalam arti sempit, istilah pekerjaan digunakan untuk suatu tugas atau kerja yang menghasilkan uang bagi seseorang.

g. Diabetes

Penderita diabetes memiliki resiko tiga kali lipat terkena stroke dan mencapai tingkat tertinggi pada usia 50-60 tahun. Setelah itu, resiko tersebut akan menurun. Namun, ada faktor penyebab lain yang dapat memperbesar resiko stroke karena sekitar 40 persen penderita diabetes pada umumnya juga mengidap hipertensi (Sulistyawati, 2013). Dalam penelitian ini, diabetes diukur melalui gula darah sesaat pasien masuk rumahsakit.

h. Asam Urat

Pengertian penyakit asam urat secara medis adalah penyakit yang diakibatkan penumpukan kristal-kristal asam urat pada persendian yang berasal dari kelebihan kadar asam urat dalam darah (AgaricPro, 2015).

i. Triglisierida

Triglisierida adalah salah satu jenis lemak utama yang mengalir di dalam darah manusia. Jika terlalu banyak, triglisierida akan menumpuk pada bagian-bagian tubuh seperti dinding pembuluh darah dan hati. Lemak jenuh dan tidak jenuh termasuk dalam kelompok triglisierida. Triglisierida sangat peting bagi manusia karena tubuh memanfaatkan lemak ini sebagai energi. Triglisierida dihasilkan tubuh dari sumber energi lainnya, yaitu karbohidrat dan dari pencernaan lemak yang terdapat dalam makanan. Tapi, kadar triglisierida yang tinggi akan meningkatkan risiko terserang penyakit serangan jantung dan stroke (Alodokter, 2015).

BAB III

METODOLOGI PENELITIAN

3.1 Sumber Data

Sumber data yang digunakan dalam penelitian ini adalah data sekunder dari data rekam medik RSUD Dr. H. Slamet Martodirdjo Pamekasan tahun 2015. Sedangkan unit penelitiannya adalah pasien penderita stroke di RSUD Dr. H. Slamet Martodirdjo Pamekasan pada tahun 2015 sebanyak 164 pasien terdiri dari 76 pasien non hemoragik dan 88 pasien hemoragik.

3.2 Variabel Penelitian

Variabel-variabel yang digunakan dalam penelitian ini adalah sebagai berikut.

a. Variabel Respon

Y (variabel respon) pada penelitian ini adalah status stroke yang diderita dengan kategori nol adalah stroke nonhemoragik (iskemik) dan satu adalah hemoragik (iskemik).

b. Variabel Prediktor

Variabel Prediktor pada penelitian ini meliputi usia pasien (X_1), jenis kelamin pasien (X_2), Hipertensi (X_3), Kolesterol (X_4), Trigliserida (X_5), Riwayat Keluarga (X_6), Asam Urat (X_7) dan Diabetes (X_8).

Tabel 3.1 Variabel Faktor-faktor yang Mempengaruhi Status Stroke

Variabel	Keterangan	Skala Data	Kategori	Definisi Operasional
Y	Status Stroke	Nominal	0 = Nonhemoragik 1 = Hemoragik	Tipe stroke pasien ditentukan ketika masuk rumahsakit
X_1	Usia Pasien	Rasio		Usia pasien tercatat dalam satuan tahun
X_2	Jenis Kelamin	Nominal	0 = Perempuan 1 = Laki-laki	Jenis kelamin pasien

Tabel 3.1 (Lanjutan)

Variabel	Keterangan	Skala Data	Kategori	Definisi Operasional
X ₃	Hipertensi	Ordinal	1 = Tidak Hipertensi (tensi darah antara 120/80 – 139/89 mmHg) 2 = Hipertensi tingkat 1 (tensi darah antara 140/90 – 159/99 mmHg) 3 = Hipertensi tingkat 2 (tensi darah antara 160/100 – 179/109 mmHg) 4 = Hipertensi tingkat 3 (tensi darah lebih dari atau sama dengan 180/110 mmHg)	Tensi darah pasien tercatat dalam satuan mmHg
X ₄	Kolesterol	Rasio		Kolesterol pasien tercatat dalam satuan mg/dl
X ₅	Trigliserida	Rasio		Trigliserida pasien tercatat dalam satuan mg/dl
X ₆	Riwayat Keluarga Stroke	Nominal	0 = Tidak Ada Riwayat Stroke 1 = Ada Riwayat Stroke	Riwayat keluarga pasien
X ₇	Asam Urat	Rasio		Kadar asam urat pasien tercatat dalam satuan mg/dl
X ₈	Diabetes	Rasio		Kadar gula darah sesaat pasien tercatat dalam satuan mg/dl

3.3 Langkah Analisis Data

Langkah-langkah dari penelitian faktor-faktor yang mempengaruhi status stroke pasien di RSUD Dr. H. Slamet Martodirdjo Pamekasan dengan menggunakan metode regresi logistik biner dan MARS adalah sebagai berikut.

1. Mengumpulkan data faktor-faktor yang mempengaruhi status stroke pasien di RSUD Dr. H. Slamet Martodirdjo Pamekasan.
2. Memasukkan data ke *software* SPSS.
3. Melakukan analisis statistika deskriptif dari data faktor-faktor yang mempengaruhi status stroke pasien di RSUD Dr. H. Slamet Martodirdjo Pamekasan.
4. Mendapatkan model MARS dengan melalui langkah sebagai berikut.
 - a. Mendefinisikan variabel respon (Y) dan variabel prediktor (X) dalam pembentukan model.
 - b. Menentukan model terbaik, dengan cara *trial and eror* sampai didapatkan model dengan nilai GCV minimum. Berikut tahapan-tahapannya.
 1. Menentukan maksimum fungsi basis (BF) yaitu dua sampai empat kali banyaknya variabel prediktor.
 2. Menentukan jumlah maksimum interaksi (MI) yaitu satu, dua dan tiga.
 3. Menentukan minimal jumlah pengamatan setiap knot atau observasi minimum (MO) antar knot sebanyak nol, satu, dua atau tiga.
 - c. Mengulangi langkah (b) sampai didapat model dengan nilai GCV minimum dan tingkat ketepatan klasifikasi terbesar.
 - d. Memperoleh model MARS berisikan variabel-variabel yang berpengaruh secara signifikan.
 - e. Menginterpretasikan model MARS terbaik.
 - f. Menghitung nilai *Sensitivity*, *Specificity* dan *Total Accuracy* dari tabel klasifikasi.
5. Mendapatkan model regresi logistik biner dengan melalui langkah sebagai berikut.

- a. Mendefinisikan variabel respon (Y) dan variabel prediktor (X) dalam pembentukan model.
 - b. Melakukan uji signifikansi parameter secara univariabel pada data faktor-faktor yang mempengaruhi status stroke pasien di RSUD Dr. H. Slamet Martodirdjo Pamekasan.
 - c. Melakukan uji signifikansi parameter secara serentak dan parsial pada data faktor-faktor yang mempengaruhi status stroke pasien di RSUD Dr. H. Slamet Martodirdjo Pamekasan.
 - d. Melakukan uji kesesuaian model dari data faktor-faktor yang mempengaruhi status stroke pasien di RSUD Dr. H. Slamet Martodirdjo Pamekasan.
 - e. Menghitung nilai *Sensitivity*, *Specificity* dan *Total Accuracy* dari tabel klasifikasi.
6. Menentukan metode yang terbaik dari kedua model tersebut untuk kasus faktor-faktor yang berpengaruh pada status stroke pasien di RSUD Dr. H. Slamet Martodirdjo Pamekasan dengan melihat nilai *Sensitivity*, *Specificity* dan *Total Accuracy* dari tabel klasifikasi.

BAB IV ANALISIS DAN PEMBAHASAN

4.1 Statistika Deskriptif Faktor-faktor yang Berpengaruh terhadap Status Stroke Pasien di RSUD Dr. H. Slamet Martodirdjo Pamekasan tahun 2015

Statistika deskriptif pada faktor-faktor yang mempengaruhi status stroke pasien di RSUD Dr. H. Slamet Martodirdjo Pamekasan pada tahun 2015 akan digambarkan pada Tabel 4.1 di bawah ini.

Tabel 4.1 Hasil Analisis Statistika Deskriptif

Nama Variabel	Kategori	Jumlah	Persentase
Status Stroke	Non Hemoragik	76	46,30%
	Hemoragik	88	53,70%
Jenis Kelamin	Perempuan	83	50,60%
	Laki-laki	81	49,40%
Hipertensi	Tidak Hipertensi	26	15,90%
	Hipertensi tingkat 1	24	14,60%
	Hipertensi tingkat 2	34	20,70%
	Hipertensi tingkat 3	80	48,80%
Riwayat Keluarga	Tidak Ada Riwayat Keluarga	145	88,40%
	Ada Riwayat Keluarga	19	11,60%
Nama Variabel	Mean	Minimum	Maximum
Usia Pasien	57	35	91
Kolesterol	195,36	49	362
Trigliserida	114,33	32	437
Asam Urat	5,08	2	20
Diabetes	144,66	28	522

Berdasarkan Tabel 4.1 dapat diketahui bahwa gambaran pasien stroke terbanyak di RSUD Dr. H. Slamet Martodirdjo Pamekasan pada tahun 2015 adalah pasien stroke hemoragik yaitu sebanyak 88 pasien dengan jenis kelamin Perempuan memiliki hipertensi tingkat tiga dengan rata-rata usia 57 tahun, rata-rata kolesterol pasien yaitu 195,36 mg/dl, trigliserida 114,33 mg/dl,

asam urat 5,08 mg/dl dan gula darah sesaat (diabetes) 144,66 mg/dl.

4.2 *Multivariate Adaptive Regression Spline (MARS)*

Pembentukan model MARS sesuai yang disarankan oleh Friedman (1991) yaitu jumlah maksimum basis fungsi (BF) dua sampai empat kali jumlah variabel prediktor, maksimum interaksi (MI) satu, dua atau tiga dengan pertimbangan jika lebih dari tiga akan menghasilkan model yang sangat kompleks dan jarak minimum antar knot (minimum observasi) sebanyak nol, satu, dua, tiga, lima atau sepuluh. Berikut hasil kombinasi dari semua BF, MI dan MO.

Tabel 4.2 Kombinasi BF, MI dan MO

BF	MI	MO	GCV	BF	MI	MO	GCV	BF	MI	MO	GCV
16	1	0	0,231	24	1	0	0,232	32	1	0	0,232
16	2	0	0,234	24	2	0	0,227	32	2	0	0,227
16	3	0	0,228	24	3	0	0,234	32	3	0	0,221
16	1	1	0,231	24	1	1	0,231	32	1	1	0,232
16	2	1	0,228	24	2	1	0,231	32	2	1	0,239
16	3	1	0,228	24	3	1	0,221	32	3	1	0,221
16	1	2	0,231	24	1	2	0,231	32	1	2	0,232
16	2	2	0,228	24	2	2	0,228	32	2	2	0,23
16	3	2	0,228	24	3	2	0,226	32	3	2	0,224
16	1	3	0,233	24	1	3	0,234	32	1	3	0,235
16	2	3	0,23	24	2	3	0,228	32	2	3	0,228
16	3	3	0,23	24	3	3	0,227	32	3	3	0,226
16	1	5	0,233	24	1	5	0,231	32	1	5	0,235
16	2	5	0,231	24	2	5	0,234	32	2	5	0,233
16	3	5	0,234	24	3	5	0,229	32	3	5	0,228
16	1	10	0,231	24	1	10	0,232	32	1	10	0,232
16	2	10	0,228	24	2	10	0,234	32	2	10	0,236
16	3	10	0,222	24	3	10	0,223	32	3	10	0,222

Menentukan model terbaik yaitu mempertimbangkan nilai GCV terkecil, bila bernilai sama maka pertimbangannya nilai MSE terkecil, bila bernilai sama pertimbangan selanjutnya yaitu ketepatan klasifikasi (*accuracy*) terbesar, bila masih belum bisa dipilih pertimbangan selanjutnya yaitu dari kombinasi BF, MI dan MO terkecil. Berdasarkan Tabel 4.2 maka dipilih model terbaik dengan GCV paling kecil yaitu sebesar 0,221 dengan nilai MSE sebesar 0,182 dengan lima variabel prediktor yang berpengaruh terhadap status stroke pasien di RSUD Dr. H. Slamet Martodirdjo Pamekasan pada tahun 2015 dimana memiliki ketepatan klasifikasi sebesar 75% yang merupakan kombinasi antara BF=24, MI=3 dan MO=1. Persamaan model MARS adalah sebagai berikut.

$$Y = 0,745 - 0,004 BF_1 - 0,005 BF_2 + 0,005 BF_3 + 0,003 BF_{12} - 0,010 BF_{20} + 0,008 BF_{22}$$

dimana,

$$BF_1 = \max(0, \text{DIABETES} - 179.000);$$

$$BF_2 = \max(0, 179.000 - \text{DIABETES});$$

$$BF_3 = (\text{HIPERTENSI} = 4) BF_1;$$

$$BF_6 = (\text{RIWAYAT KELUARGA} = 1);$$

$$BF_8 = (\text{HIPERTENSI} = 1 \text{ OR HIPERTENSI} = 4);$$

$$BF_{12} = \max(0, 157.000 - \text{TRIGLISERIDA}) BF_8;$$

$$BF_{20} = \max(0, 132.000 - \text{TRIGLISERIDA}) BF_6;$$

$$BF_{22} = \max(0, 50.000 - \text{USIA PASIEN}) BF_{20};$$

Dapat diketahui bahwa model yang terbaik adalah model dengan Maksimum Interaksi sebanyak tiga maka terdapat beberapa variabel yang saling berinteraksi dalam mempengaruhi status stroke pasien di RSUD Dr. H. Slamet Martodirdjo Pamekasan.

Untuk melihat faktor apa saja yang mempengaruhi status stroke pasien di RSUD Dr. H. Slamet Martodirdjo Pamekasan pada tahun 2015 dapat dilihat melalui interpretasi *Odds ratio*. *Odds ratio* adalah suatu ukuran yang menunjukkan rasio untuk mengalami suatu kejadian tertentu antara suatu bagian populasi

dengan ciri tertentu dan bagian populasi lain yang tidak memiliki ciri tertentu.

$$OR = \frac{e^{(0,745-0,004 BF_1-0,005 BF_2+0,005 BF_3+0,003 BF_{12}-0,01 BF_{20}+0,008 BF_{22})}}{1+e^{(0,745-0,004 BF_1-0,005 BF_2+0,005 BF_3+0,003 BF_{12}-0,01 BF_{20}+0,008 BF_{22})}}$$

$$OR BF_1 = \frac{e^{(0,745-0,004 (179))}}{1+e^{(0,745-0,004 (179))}} = 0,5071$$

$$OR BF_2 = \frac{e^{(0,745-0,005 (179))}}{1+e^{(0,745-0,005 (179))}} = 0,4625$$

$$OR BF_3 = \frac{e^{(0,745+0,005 (4))}}{1+e^{(0,745+0,005 (4))}} = 0,6825$$

$$OR BF_{12} = \frac{e^{(0,745+0,003 (157))}}{1+e^{(0,745+0,003 (157))}} = 0,7713$$

$$OR BF_{20} = \frac{e^{(0,745-0,010 (132))}}{1+e^{(0,745-0,010 (132))}} = 0,3601$$

$$OR BF_{22} = \frac{e^{(0,745+0,008 (50))}}{1+e^{(0,745+0,008 (50))}} = 0,7586$$

Tabel 4.3 Nilai *Odds Ratio* (OR) untuk model MARS

No	BF	Koefisien	OR
1	BF ₁	-0,004	0,5071
2	BF ₂	-0,005	0,4625
3	BF ₃	0,005	0,6825
4	BF ₁₂	0,003	0,7713
5	BF ₂₀	-0,01	0,3601
6	BF ₂₂	0,008	0,7586

Interpretasi model berdasarkan persamaan yang terbentuk pada model MARS di atas adalah sebagai berikut.

- Koefisien BF₁ = max(0, DIABETES - 179.000) dengan *odds ratio* sebesar 0,5071 dapat diartikan bahwa jika kadar gula darah sesaat (diabetes) pasien lebih dari 179 mg/dl maka peluang pasien terkena stroke hemoragik yaitu 0,5071 dibandingkan terkena stroke non hemoragik.
- Koefisien BF₂ = max(0, 179.000 - DIABETES) dengan *odds ratio* sebesar 0,4625 dapat diartikan bahwa jika kadar

- gula darah sesaat (diabetes) pasien kurang dari 179 mg/dl maka peluang pasien terkena stroke hemoragik yaitu 0,4625 dibandingkan terkena stroke non hemoragik.
- Koefisien $BF_3 = (\text{HIPERTENSI} = 4) BF_1$
 $BF_1 = \max(0, \text{DIABETES} - 179.000)$ dengan *odds ratio* sebesar 0,6825 dapat diartikan bahwa jika kadar gula darah sesaat (diabetes) pasien lebih dari 179 mg/dl dan hipertensi tingkat tiga ($\geq 180/110$ mmHg) maka peluang pasien terkena stroke hemoragik yaitu 0,6825 dibandingkan terkena stroke non hemoragik.
 - Koefisien $BF_{12} = \max(0, 157.000 - \text{TRIGLISERIDA}) BF_8$;
 $BF_8 = (\text{HIPERTENSI} = 1 \text{ OR } \text{HIPERTENSI} = 4)$ dengan *odds ratio* sebesar 0,7713 dapat diartikan bahwa jika kadar trigliserida pasien kurang dari 157 mg/dl dan tidak hipertensi atau hipertensi tingkat tiga ($\geq 180/110$ mmHg) maka peluang pasien terkena stroke hemoragik yaitu 0,7713 dibandingkan terkena stroke non hemoragik.
 - Koefisien $BF_{20} = \max(0, 132.000 - \text{TRIGLISERIDA}) BF_6$;
 $BF_6 = (\text{RIWAYAT KELUARGA}=1)$ dengan *odds ratio* sebesar 0,3601 dapat diartikan bahwa jika kadar trigliserida pasien kurang dari 132 mg/dl dan memiliki riwayat keluarga stroke maka kecenderungan pasien terkena stroke hemoragik yaitu 0,3601 dibandingkan terkena stroke non hemoragik.
 - Koefisien $BF_{22} = \max(0, 50.000 - \text{USIA PASIEN}) BF_{20}$;
 $BF_{20} = \max(0, 132.000 - \text{TRIGLISERIDA}) * BF_6$ dengan *odds ratio* sebesar 0,7586 dapat diartikan bahwa jika usia pasien kurang dari 50 tahun dan memiliki kadar trigliserida lebih dari 132 mg/dl maka peluang pasien terkena stroke hemoragik 0,7586 dibandingkan terkena stroke non hemoragik.

Tingkat kepentingan dari setiap variabel ditaksir oleh kenaikan nilai GCV. Dengan demikian tingkat kepentingan setiap variabel mempunyai peranan untuk meminimumkan nilai GCV di dalam model. Adapun tingkat kepentingan dari masing-masing

variabel tersebut terhadap model tersaji pada Tabel 4.4 di bawah ini.

Tabel 4.4 Tingkat Kepentingan Variabel dalam MARS

Variabel	Importance	Pengurangan GCV
Diabetes	100,000	0,246
Hipertensi	83,546	0,239
Trigliserida	67,497	0,233
Riwayat Keluarga	62,959	0,231
Usia Pasien	40,133	0,225
Jenis Kelamin	0,000	0,221
Kolesterol	0,000	0,221
Asam Urat	0,000	0,221

Pada Tabel 4.4 dapat diketahui bahwa variabel diabetes merupakan variabel terpenting pada model MARS status stroke pasien di RSUD Dr. H. Slamet Martodirdjo Pamekasan pada tahun 2015 dengan tingkat kepentingan mencapai 100%. Selanjutnya variabel hipertensi mencapai 83,546%, trigliserida 67,497%, riwayat keluarga 62,959% dan usia pasien 40,133%. Nilai minus GCV menunjukkan bahwa apabila suatu variabel itu dikeluarkan dari model maka nilai GCV akan menambah sebesar +GCV pada variabel tersebut. Sehingga semakin besar tingkat kepentingan dari variabel tersebut akan diikuti nilai minus GCV yang besar dan akan meminimumkan nilai GCV dalam model yang terbentuk.

4.3 Regresi Logistik Biner

4.3.1 Uji Signifikansi Parameter Secara Univariabel

Uji signifikansi parameter secara univariabel digunakan untuk mengetahui faktor-faktor yang mempengaruhi status stroke pasien di RSUD Dr. H. Slamet Martodirdjo Pamekasan tahun

2015. Berikut hasil uji univariabel terhadap status stroke pasien di RSUD Dr. H. Slamet Martodirdjo Pamekasan.

Hipotesis :

$$H_0: \beta_j = 0$$

$$H_1: \beta_j \neq 0; j = 1, 2, \dots, 5.$$

Tabel 4.5 Uji Signifikansi Parameter Secara Univariabel

Variabel	Koefisien (B)	SE	Wald	df	P-value	Exp (B)
Diabetes	0,005	0,003	3,463	1	0,063	1,005
<i>constant</i>	-0,518	0,384	1,824	1	0,177	0,595
Hipertensi			12,333	3	0,006	3,927
Hipertensi (1)	-0,885	0,46	3,699	1	0,054	0,413
Hipertensi (2)	-1,067	0,478	4,988	1	0,026	0,344
Hipertensi (3)	-1,337	0,431	9,623	1	0,002	0,263
<i>constant</i>	0,731	0,239	9,375	1	0,002	2,077
Trigliserida	0,00	0,003	0,036	1	0,849	1,000
<i>constant</i>	0,204	0,339	0,362	1	0,547	1,226
Riwayat Keluarga(1)	1,037	0,521	3,957	1	0,047	2,820
<i>constant</i>	-0,773	0,494	2,454	1	0,117	0,462
Usia Pasien	-0,006	0,014	0,202	1	0,653	0,994
<i>constant</i>	0,505	0,813	0,386	1	0,534	1,658

Berdasarkan Tabel 4.5 dapat diketahui bahwa terdapat variabel yang signifikan berpengaruh terhadap status stroke pasien di RSUD Dr. H. Slamet Martodirdjo Pamekasan yaitu pasien yang memiliki hipertensi tingkat satu, hipertensi tingkat dua dan memiliki riwayat keluarga stroke.

4.3.2 Uji Signifikansi Parameter Secara Multivariabel

Uji signifikan serentak digunakan untuk memperoleh model yang tepat dan sederhana berdasarkan faktor-faktor yang dianggap berpengaruh terhadap variabel respon. Dimana faktor-faktor yang berpengaruh terhadap status stroke pasien di RSUD Dr. H. Slamet Martodirdjo Pamekasan akan dianalisis secara serentak. Berikut hasil uji signifikansi parameter secara serentak.

Hipotesis :

$$H_0: \beta_1 = \beta_2 = \beta_3 = 0$$

H_1 : Paling tidak terdapat satu $\beta_j \neq 0; j = 1, 2, 3$.

Tabel 4.6 Uji Signifikansi Parameter Secara Serentak

	<i>Chi-square</i>	<i>df</i>	<i>Pvalue</i>
<i>Step</i>	17,184	4	0,002
<i>Block</i>	17,184	4	0,002
<i>Model</i>	17,184	4	0,002

Berdasarkan Tabel 4.6 dapat diketahui bahwa *Pvalue* yang dihasilkan sebesar 0,002 sehingga diperoleh keputusan yaitu tolak H_0 maka dapat memberikan informasi bahwa terdapat cukup bukti bahwa paling tidak terdapat salah satu variabel prediktor yang berpengaruh signifikan terhadap status penyakit stroke secara serentak.

Uji signifikansi parameter secara parsial digunakan untuk mengetahui faktor-faktor yang mempengaruhi status stroke pasien di RSUD Dr. H. Slamet Martodirdjo Pamekasan secara parsial. Berikut hasil uji parsial terhadap status stroke pasien di RSUD Dr. H. Slamet Martodirdjo Pamekasan.

Hipotesis :

$$H_0: \beta_j = 0$$

H_1 : $\beta_j \neq 0; j = 1, 2, 3$.

Tabel 4.7 Uji Signifikansi Parameter Secara Parsial

	<i>B</i>	<i>S.E.</i>	<i>Wald</i>	<i>df</i>	<i>Sig.</i>	<i>Exp(B)</i>
Hipertensi			12,283	3	0,006	0,996
Hipertensi (1)	-0,814	0,469	3,009	1	0,083	0,443
Hipertensi (2)	-1,070	0,484	4,880	1	0,027	0,343
Hipertensi (3)	-1,379	0,436	9,990	1	0,002	0,252
Riwayat Keluarga (1)	1,077	0,541	3,966	1	0,046	2,937
<i>Constant</i>	-0,227	0,536	0,179	1	0,672	0,797

Berdasarkan Tabel 4.7 dapat diketahui bahwa variabel yang signifikan terhadap status stroke pasien di RSUD Dr. H. Slamet Martodirdjo pada tahun 2015 adalah hipertensi tingkat satu,

hipertensi tingkat dua dan memiliki riwayat keluarga sehingga diperoleh model yaitu

$$\hat{\pi}(x) = -0,227 - 0,814 \text{ Hipertensi}(1) - 1,070 \text{ Hipertensi}(2) - 1,379 \text{ Hipertensi}(3) + 1,077 \text{ Riwayat Keluarga (1)}$$

Disamping itu untuk menginterpretasikan model dapat dilihat dari nilai $\text{Exp}(B)$ atau *Odds Ratio* pada Tabel 4.7. Dari nilai *Odds Ratio* dapat diambil kesimpulan sebagai berikut.

1. Variabel Hipertensi (1) : Jika pasien tidak hipertensi maka peluang terkena stroke hemoragik yaitu 0,443 dibandingkan pasien yang memiliki hipertensi tingkat tiga (tensi darah lebih dari atau sama dengan 180/110 mmHg).
2. Variabel Hipertensi (2) : Jika pasien memiliki hipertensi tingkat satu (tensi darah antara 140/90 mmHg sampai 159/99 mmHg) maka peluang terkena stroke hemoragik yaitu 0,343 dibandingkan pasien yang memiliki hipertensi tingkat tiga (tensi darah lebih dari atau sama dengan 180/110 mmHg).
3. Variabel Hipertensi (3) : Jika pasien memiliki hipertensi tingkat dua (tensi darah lebih dari 160/100 mmHg sampai 179/109 mmHg) maka peluang terkena stroke hemoragik yaitu 0,252 dibandingkan pasien yang memiliki hipertensi tingkat tiga (tensi darah lebih dari atau sama dengan 180/110 mmHg).
4. Variabel Riwayat Keluarga (1) : Jika pasien memiliki riwayat keluarga stroke maka kecenderungan terkena stroke hemoragik 2,937 kali lipat dibandingkan pasien yang tidak memiliki riwayat keluarga stroke.

4.3.3 Kesesuaian Model

Untuk menguji kesesuaian model apakah model sesuai dalam artian tidak ada perbedaan antara hasil observasi dengan kemungkinan hasil prediksi model dilakukan dengan menggunakan hipotesis sebagai berikut.

H_0 : Model sesuai (ada hubungan antara hasil pengamatan dengan hasil prediksi).

H_1 : Model tidak sesuai (tidak ada hubungan antara hasil pengamatan dengan hasil prediksi).

Berdasarkan hasil uji kesesuaian model berdasarkan Lampiran B dapat diketahui bahwa *Pvalue* yang dihasilkan adalah 0,610 sehingga menghasilkan keputusan gagal tolak H_0 maka model sesuai (ada hubungan antara hasil pengamatan dengan hasil prediksi).

4.4 Ketepatan Klasifikasi pada Model

Ketepatan klasifikasi digunakan untuk mengetahui seberapa besar peluang ketepatan model dalam pengklasifikasian status stroke pasien di RSUD Dr. H. Slamet Martodirdjo Pamekasan dengan menghitung nilai *Sensitivity*, *Specificity* dan *Total Accuracy* dari tabel hasil klasifikasi.

4.4.1 Ketepatan Klasifikasi pada Model Regresi Logistik Biner

Berikut hasil ketepatan klasifikasi dengan menggunakan metode regresi logistik biner mengenai faktor-faktor yang mempengaruhi status stroke pasien di RSUD Dr. H. Slamet Martodirdjo Pamekasan pada tahun 2015.

Tabel 4.8 Hasil Klasifikasi Model Regresi Logistik Biner

Kelompok Observasi	Status Stroke		Total
	Non Hemoragik	Hemoragik	
Non Hemoragik	44	32	76
Hemoragik	27	61	88
Total	71	93	164

Tabel 4.9 Hasil Klasifikasi Model MARS

Kelompok Observasi	Status Stroke		Total
	Non Hemoragik	Hemoragik	
Non Hemoragik	48	28	76
Hemoragik	13	75	88
Total	61	103	164

Berdasarkan Tabel 4.9 dan 4.10 maka dapat dihitung nilai nilai *Sensitivity*, *Specificity* dan *Total Accuracy* dari kedua tabel hasil klasifikasi tersebut.

Tabel 4.10 Hasil Perhitungan Nilai *Sensitivity*, *Specificity* dan *Total Accuracy*

Metode	<i>Sensitivity</i>	<i>Specificity</i>	<i>Total Accuracy</i>
Regresi Logistik Biner	0,693	0,579	0,640
MARS	0,852	0,632	0,750

Perbandingan hasil klasifikasi antara kedua model mengenai faktor yang mempengaruhi status stroke pasien di RSUD Dr. H. Slamet Martodirdjo Pamekasan pada tahun 2015 dapat diketahui bahwa model MARS memiliki nilai *Total Accuracy* paling tinggi dibandingkan dengan model regresi logistik biner. Dimana *Total Accuracy* merupakan nilai persentase ketepatan klasifikasi secara keseluruhan, maka klasifikasi yang dihasilkan adalah semakin baik.

LAMPIRAN**A. Data Hasil Rekam Medik Mengenai Pasien Stroke di RSUD
Dr. H. Slamet Martodirdjo Pamekasan tahun 2015**

No	status stroke	usia pasien	jenis kelamin	hiper- tensi	koles- terol	trigliser- ida	riwayat keluarga	asma murat	diabete- s
1	infark	35	P	2	269	170	1	10	88
2	infark	42	L	1	202	108	0	5,3	97
3	infark	74	L	2	175	152	0	4	101
4	infark	59	P	2	240	82	0	5,5	93
5	infark	50	P	1	172	106	0	4,2	137
6	infark	67	P	4	216	112	0	2,3	60
7	infark	70	L	3	195	68	0	5,9	97
8	infark	70	P	4	229	128	0	6,1	106
9	bleeding	55	L	1	160	97	0	4,7	109
10	infark	65	P	3	269	437	0	2,2	424
11	infark	52	L	2	128	54	0	2,6	91
12	infark	75	P	1	128	161	0	7	95
13	infark	59	L	2	151	76	0	2,8	98

14	infark	57	P	4	210	124	0	10,2	79
15	infark	63	P	3	139	74	0	2,6	107
16	infark	65	P	1	219	402	0	3	460
17	bleedin g	54	L	1	164	118	0	2,7	84
18	bleedin g	40	P	1	212	119	0	3,6	200
19	bleedin g	70	P	3	258	111	0	5,9	140
:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:
16 0	bleedin g	37	P	1	194	83	0	2	98
16 1	bleedin g	40	P	4	186	121	0	2,5	119
16 2	bleedin g	48	P	4	190	62	1	4,5	135
16 3	bleedin g	67	P	4	184	145	0	3,1	144
16 4	bleedin g	61	L	2	214	118	0	5,9	176

Lampiran B. Hasil Output SPSS

1. Hasil *Output* SPSS Statistika Deskriptif

statusstroke

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	non hemoragik	76	46.3	46.3	46.3
	hemoragik	88	53.7	53.7	100.0
	Total	164	100.0	100.0	

jeniskelamin

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Perempuan	83	50.6	50.6	50.6
	Laki-laki	81	49.4	49.4	100.0
	Total	164	100.0	100.0	

hipertensi

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	tidak hipertensi	26	15.9	15.9	15.9

hipertensi tingkat 1	24	14.6	14.6	30.5
hipertensi tingkat 2	34	20.7	20.7	51.2
hipertensi tingkat 3	80	48.8	48.8	100.0
Total	164	100.0	100.0	

riwayatkeluarga

	Frequency	Percent	Valid Percent	Cumulative Percent
Tidak Ada Riwayat Keluarga	145	88.4	88.4	88.4
Valid Ada Riwayat Keluarga	19	11.6	11.6	100.0
Total	164	100.0	100.0	

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
usiap pasien	164	35.00	91.00	57.3110	11.29502
kolesterol	164	49.00	362.00	195.3659	51.20667
trigliserida	164	32.00	437.00	114.3354	59.79692
asamurat	164	2.00	20.00	5.0811	2.65904
diabetes	164	28.00	522.00	144.6585	73.01336
Valid N (listwise)	164				

2. *Output* SPSS untuk Analisis Regresi Logistik Biner
 Pengujian Secara Univariabel

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a diabetes	.005	.003	3.463	1	.063	1.005
Constant	-.518	.384	1.824	1	.177	.595

a. Variable(s) entered on step 1: diabetes.

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a hipertensi			12.333	3	.006	
hipertensi(1)	-.885	.460	3.699	1	.054	.413
hipertensi(2)	-1.067	.478	4.988	1	.026	.344
hipertensi(3)	-1.337	.431	9.623	1	.002	.263
Constant	.731	.239	9.375	1	.002	2.077

a. Variable(s) entered on step 1: hipertensi.

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a triglicerida	.000	.003	.036	1	.849	1.000
Constant	.204	.339	.362	1	.547	1.226

a. Variable(s) entered on step 1: triglicerida.

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
riwayatkeluarga(1)	1.037	.521	3.957	1	.047	2.820
Step 1 ^a						
Constant	-.773	.494	2.454	1	.117	.462

a. Variable(s) entered on step 1: riwayatkeluarga.

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a usiapasien	-.006	.014	.202	1	.653	.994
Constant	.505	.813	.386	1	.534	1.658

a. Variable(s) entered on step 1: usiapasien.

3. Pengujian Secara Seentak dan Parsial

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	17.184	4	.002
	Block	17.184	4	.002
	Model	17.184	4	.002

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a			12.283	3	.006	
hipertensi						
hipertensi(1)	-.814	.469	3.009	1	.083	.443
hipertensi(2)	-1.070	.484	4.880	1	.027	.343
hipertensi(3)	-1.379	.436	9.990	1	.002	.252
riwayatkeluarga(1)	1.077	.541	3.966	1	.046	2.937
Constant	-.227	.536	.179	1	.672	.797

a. Variable(s) entered on step 1: hipertensi, riwayatkeluarga.

4. Uji Kesesuaian Model

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	2.696	4	.610

5. Tabel Klasifikasi

Classification Table^a

Observed		Predicted		
		statusstroke		Percentage Correct
		non hemoragik	hemoragik	
Step 1	non hemoragik	44	32	57.9
	hemoragik	27	61	69.3
	Overall Percentage			64.0

a. The cut value is ,500

Lampiran C. *Output* MARS 2.0

VARIABLES IN RECT FILE ARE:
 STATUSST USIAPASI JENISKEL HIPERTEN KOLESTER
 TRIGLISE RIWAYATK ASAMURAT DIABETES

C:\MARS 2.0\SPSS TA Dhinta.SAV[spsswin]: 164 RECORDS.

MARS VERSION 2.0.0.19

READING DATA, UP TO 1996351 RECORDS.

RECORDS READ: 164

RECORDS KEPT IN LEARNING SAMPLE: 164

LEARNING SAMPLE STATISTICS

=====

VARIABLE	MEAN	SD	N	SUM
-----	-----	-----	-----	-----
STATUSST	0.537	0.500	164.000	88.000
USIAPASI	57.311	11.295	164.000	9399.000
JENISKEL	0.494	0.501	164.000	81.000
HIPERTEN	3.024	1.129	164.000	496.000
KOLESTER	195.366	51.207	164.000	32040.000
TRIGLISE	114.335	59.797	164.000	18751.000
RIWAYATK	0.116	0.321	164.000	19.000
ASAMURAT	5.081	2.659	164.000	833.300
DIABETES	144.659	73.013	164.000	23724.000

Final Model (After Backward Stepwise Elimination)

=====

Basis Fun	Coefficient	Variable	Parent	Knot
-----	-----	-----	-----	-----
0	0.745			
1	-0.004	DIABETES		179.000
2	-0.005	DIABETES		179.000
3	0.005	HIPERTEN	DIABETES	
12	0.003	TRIGLISE	HIPERTEN	157.000
20	-0.010	TRIGLISE	RIWAYATK	132.000
22	0.008	USIAPASI	TRIGLISE	50.000

Piecewise Linear GCV = 0.221, #efprms = 18.368

Relative Variable Importance

=====

Variable	Importance	-gcv
-----	-----	-----
8 DIABETES	100.000	0.246
3 HIPERTEN	83.546	0.239
5 TRIGLISE	67.497	0.233
6 RIWAYATK	62.959	0.231
1 USIAPASI	40.133	0.225
2 JENISKEL	0.000	0.221
4 KOLESTER	0.000	0.221
7 ASAMURAT	0.000	0.221

ORDINARY LEAST SQUARES RESULTS
=====

N: 164.000 R-SQUARED: 0.298
MEAN DEP VAR: 0.537 ADJ R-SQUARED: 0.271
UNCENTERED R-SQUARED = R-0 SQUARED: 0.675

PARAMETER	ESTIMATE	S.E.	T-RATIO	P-VALUE
Constant	0.745	0.075	9.941	.999201E-15
Basis Function 1	-0.004	.991813E-03	-3.681	.318887E-03
Basis Function 2	-0.005	0.001	-4.894	.242629E-05
Basis Function 3	0.005	0.001	3.667	.335165E-03
Basis Function 12	0.003	.921037E-03	3.413	.816865E-03
Basis Function 20	-0.010	0.002	-4.224	.405668E-04
Basis Function 22	0.008	0.003	3.041	0.003

F-STATISTIC =	11.100	S.E. OF REGRESSION =		0.427
P-VALUE =	.264988E-09	RESIDUAL SUM OF SQUARES =		28.634
[MDF,NDF] = [6, 157]		REGRESSION SUM OF SQUARES =		12.147

The Following Graphics Are Piecewise Linear

Basis Functions
=====

```
BF1 = max(0, DIABETES - 179.000);
BF2 = max(0, 179.000 - DIABETES );
BF3 = (HIPERTEN = 4) * BF1;
BF6 = (RIWAYATK = 1);
BF8 = (HIPERTEN = 1 OR HIPERTEN = 4);
BF12 = max(0, 157.000 - TRIGLISE ) * BF8;
BF20 = max(0, 132.000 - TRIGLISE ) * BF6;
BF22 = max(0, 50.000 - USIAPASI ) * BF20;

Y = 0.745 - 0.004 * BF1 - 0.005 * BF2 + 0.005 * BF3 + 0.003 * BF12
  - 0.010 * BF20 + 0.008 * BF22;

model STATUSST = BF1 BF2 BF3 BF12 BF20 BF22;
```

=====
LEARNING SAMPLE CLASSIFICATION TABLE
=====

Actual Class	Predicted Class		Actual Total
	0	1	
0	48.000	28.000	76.000
1	13.000	75.000	88.000

Pred. Tot.	61.000	103.000	164.000
Correct	0.632	0.852	
Success Ind.	0.168	0.316	
Tot. Correct	0.750		

Sensitivity: 0.632 Specificity: 0.852
False Reference: 0.213 False Response: 0.272
Reference = Class 0, Response = Class 1

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan analisis dan pembahasan yang telah diuraikan sebelumnya maka dapat diambil kesimpulan sebagai berikut.

1. Berdasarkan hasil analisis menggunakan metode MARS diperoleh model yaitu $Y = 0,745 - 0,004 BF_1 - 0,005 BF_2 + 0,005 BF_3 + 0,003 BF_{12} - 0,010 BF_{20} + 0,008 BF_{22}$ dimana faktor-faktor yang berpengaruh terhadap status stroke pasien di RSUD Dr. H. Slamet Martodirdjo Pamekasan adalah diabetes, hipertensi, riwayat keluarga stroke, trigliserida dan usia pasien. Sedangkan berdasarkan hasil metode regresi logistik biner diperoleh model $\hat{\pi}(x) = -0,227 - 0,814 \text{ Hipertensi}(2) - 1,070 \text{ Hipertensi}(3) + 1,077 \text{ Riwayat Keluarga}(1)$ dimana faktor-faktor yang mempengaruhi status stroke pasien di RSUD Dr. H. Slamet Martodirdjo Pamekasan yaitu hipertensi tingkat satu, hipertensi tingkat dua dan memiliki riwayat keluarga.
2. Berdasarkan hasil ketepatan klasifikasi antara model regresi logistik biner dan MARS pada status stroke pasien di RSUD Dr. H. Slamet Martodirdjo Pamekasan pada tahun 2015 diperoleh kesimpulan bahwa model MARS memiliki nilai *Total Accuracy* paling tinggi dibandingkan dengan model regresi logistik biner. Sehingga pemodelan faktor yang berpengaruh terhadap status stroke pasien di RSUD Dr. H. Slamet Martodirdjo Pamekasan pada tahun 2015 lebih tepat menggunakan metode MARS.

5.2 Saran

Saran yang dapat diberikan dalam penelitian ini adalah sebagai berikut.

1. Masyarakat lebih baik menghindari makanan yang memiliki kadar gula tinggi dan berolahraga yang cukup untuk menjaga kondisi tubuh tetap stabil dengan bertambahnya usia.
2. Metode MARS merupakan metode yang baik dalam prediksi, tetapi dibutuhkan ketelitian yang cukup dalam melakukan kombinasi antara BF, MI dan MO sehingga hasil yang diperoleh tidak salah.

DAFTAR PUSTAKA

- Agresti, A. (2002). *Categorical Data Analysis*, John Wiley & Sonc, Inc, New.
- Alodokter. (2015). Gejala, Penyebab dan Mengobati. Diakses kembali pada 16 Maret 2016 dari <http://www.alodokter.com/triglicerida>.
- ArgaricPro. (2015). Pengertian Penyakit Asam Urat. Diakses kembali pada 04 November 2015, dari <http://www.agaricpro.com/-pengertian-penyakit-asam-urat/>.
- Batticaca, F. B. (2012). *Asuhan Keperawatan pada Klien dengan Gangguan Sistem Persarafan*. Jakarta:Salemba Medika.
- Bishop, Y.M., Flenberg, S.E., dan Holland, P.W. (2007). *Discreate Multivariate Analysis : Theory and Practice*, Reprint of The 1975 edition, Springer Science + Business Media, New York.
- Cherkassky, V., & Mulier, F. M. (1998). *Learning from Data: Concepts, Theory, and Methods*. John Wiley & Sons.
- Christanto, R., Mahama N.C., dan Tumboimbela, J.M. (2014). Profil Faktor-Faktor Risiko pada Pasien Stroke yang Dirawat Inap di Irina F Neurologi RSUP Prof. Dr. R. D. Kandou Manado Periode Januari – Oktober 2012. *Jurnal e-Clinic (eCi)*, Vol. 2, No. 3, Fakultas Kedokteran, Universitas Sam Ratulangi Manado.
- Dinas Kesehatan. (2012). 10 Penyakit Terbanyak di Wilayah Kerja Dinas Kesehatan Kabupaten Pamekasan Tahun 2012. Diakses kembali pada 10 Februari 2016, dari <http://dinkes.pamekasan-kab.go.id/index.php/faq/98artist72205915/art79216934/164-1996002194>.
- Gaul. (2013). Penyebab Penyakit Stroke ,Gejala dan Solusi Pencegahan. Diakses kembali pada 04 November 2015, dari <http://www.doktergaul.net/-2013/08/Gejala-Penyabab-dan-Pencegahan-Penyakit-Stroke-.html>.
- Friedman, J. H. (1991). Multivariate Adaptive Regression Splines. *The Annals of Statistics*, Vol.19, No. 1.
- Futurebeauty. (2012). Pentingnya Melakukan Tes Kadar Kolesterol Darah. Diakses kembali pada 04 November 2015, dari <http://artikeltentang-kesehatan.com/pentingnyamelakukan-test-kadar-kolesterol-darah.html>.
- Hosmer, D.W. dan Lemeshow, S. (2000). *Applied Logistic Regression*. John Wiley and Sons (2th ed.). USA.
- Kabi, G.R.C.Y, Tumewah, R. dan Kembuan, M.A.H.N. (2015). Gambaran Faktor Risiko pada Penderita Stroke Iskemik yang Dirawat Inap

- Neurologi RSUP Prof. Dr. R. D. Kandou Manado Periode Juli 2012 – Juni 2013. Jurnal *e-Clinic* (eCi), Vol. 3, No. 1, Fakultas Kedokteran, Universitas Sam Ratulangi Manado.
- Kompas. (2011). 9 Penyakit yang Diwariskan Keluarga. Diakses kembali pada 10 Februari 2016 dari <http://health.kompas.com/read/2011/09/07/10340378/9.Penyakit.yang.Diwariskan.Keluarga>.
- Muhlisin, A. (2015). Faktor Resiko Stroke. Diakses kembali pada 04 November 2015 dari <http://www.medicus.com>.
- Murti, Bhisma. (2008). *Prinsip dan Metode Riset Epidemiologi*. Gadjah Mada University Press, Yogyakarta.
- Muttaqin, A. (2008). *Asuhan Keperawatan Klien dengan Gangguan Sistem Persarafan*. Jakarta:Salemba Medika.
- Nash, M. S., & Bradford, D. F. (2001). *Parametric and Nonparametric Logistic Regressions for Prediction of Presence/Absence of an Amphibian*, Las Vegas, Nevada.
- Oktafia, D. S. (2013). Analisis Terhadap Faktor-faktor yang Mempengaruhi Laju Perbaikan Kondisi Klinis Pasien Penderita Stroke dengan Regresi Cox Weibull. Jurnal Sains dan Seni POMITS Vol. 2, No.1. FMIPA, Statistika: ITS.
- Setiawan, B. (2014). Pengertian Pekerjaan Profesi dan Professional. Diakses kembali pada 04 November 2015 dari <http://www.seputarpendidikan.com/2014/08/pengertianpekerjaan-profesi-dan.html>.
- Sulistiyawati, D. (2013). Apa Penyebab Stroke. Diakses kembali pada 26 Januari 2014, dari <http://dinkeskebumen.wordpress.com/2013/04/18/apa-penyebab-stroke/>.
- Sumarno. (2014). Model Klasifikasi Menggunakan Metode Regresi Logistik dan *Multivariate Adaptive Regression Splines* (MARS) (Studi Kasus: Data Survei Biaya Hidup (SBH) Kota Kediri Tahun 2012). Tesis. Program Magister Jurusan Statistika: Institut Teknologi Sepuluh Nopember.
- Surya, D. W. T. S. (2014). Faktor-Faktor yang Mempengaruhi Status Penyakit Stroke Pasien di RSUD Dr. H. Slamet Martodirdjo Pamekasan dengan Menggunakan Metode Regresi Logistik Biner. Tugas Akhir. Diploma Tiga Statistika: Institut Teknologi Sepuluh Nopember.
- Widowati, U. (2015). 10 Penyakit Paling Mematikan di Indonesia. Diakses kembali pada 06 Februari 2016, dari <http://www.cnnindonesia.com/gaya-hidup/20150513163407-255-53129/10-penyakit-paling-mematikan-di-indonesia/>.

- Yasril dan Kasjono. (2008). *Analisis Multivariat*. Mitra Cendekia Press, Yogyakarta.
- Zambrano, O., Rocco, S., Claudio, M., dan Museli, M. (2005). *Estimating Female Labor Force Participation Through Statistical and Machine Learning Methods: A Comparason, Banco Central de Venezuela and Kennedy School of Goverment MPA/ID*.

BIODATA PENULIS



Penulis terlahir di Pamekasan, pada tanggal 5 Januari 1994 yang bernama lengkap Dhinta Wulansari Tri Surya sebagai anak ketiga dari lima bersaudara. Penulis menyelesaikan pendidikan formal di TK Bhayangkari, SDN Lawangan Daya II, SMPN 2 Pamekasan, dan SMAN 2 Pamekasan. Selanjutnya pada tahun 2011 penulis di terima di ITS tepatnya pada program diploma DIII Statistika dan pada tahun 2014 melanjutkan studi melalui program Lintas Jalur S1 Statistika dan terdaftar dengan NRP 1314105026.

Penulis masuk ke Diploma Statistika melalui jalur ujian masuk program DIII. Penulis masuk di jurusan statistika ITS sebagai angkatan ke 22 ($\Sigma 22.060$). Penulis pernah menjadi PJ kota Pamekasan dalam satu acara STATION (*Statistics Competition*). Penulis mengerjakan Tugas Akhir di bawah naungan Laboratorium Lingkungan dan Kesehatan dengan dosen pembimbing ibu Ir. Sri Pingit Wulandari, M.Si dan dosen co. pembimbing bapak Dr. Bambang Widjanarko Otok, S.Si, M.Si. Penulis mengambil topik tentang faktor-faktor yang mempengaruhi status penyakit stroke dengan perbandingan ketepatan klasifikasi metode regresi logistik biner dan MARS. Bagi pembaca yang mempunyai saran dan kritik yang membangun mengenai tulisan ini bisa disampaikan atau didiskusikan melalui e-mail : dhintasayangmama@gmail.com dan nomer hp : 082301877320.